COUPLING LIMIT ORDER BOOKS AND BRANCHING RANDOM WALKS

FLORIAN SIMATOS

Eindhoven University of Technology

Abstract. We consider a model for a one-sided limit order book proposed by Lakner et al. [9] and show that it can be coupled with a branching random walk. We then use the coupling to answer non-trivial questions about the long-term behavior of the price. The coupling relies on a classical idea of enriching the state-space by artificially creating a filiation, in this context between orders of the book, that we believe has the potential of being useful for a broader class of models.

1. Introduction

Limit order books. A limit order book is a financial trading mechanism that keeps track of orders made by traders, and allows to execute them in the future. Typically, a trader places an order to buy a commodity at a certain level x. If the current price of the commodity is larger than x when the order is placed, then the order is kept in the book and may be fulfilled later in the future, as the price of the commodity fluctuates and falls below x. Similarly, traders may place sell orders, which gives rise to two-sided order books. Because of the importance of limit order books in financial markets, there has been a lot of research on these models, see for instance the survey by Gould et al. [5].

There are many variants of limit order books, depending for instance on which information of the book traders have access to. Typically, traders have only access to the current so-called bid and ask prices, that correspond to the lowest sell order and the highest buy order, respectively. This gives an incentive for traders to place orders in the vicinity of these prices, thus creating an intricate dynamic where the state of the book influences its evolution. Stochastic models capturing this dynamic have for instance been proposed in Cont et al. [4], Lakner et al. [9] and Yudovina [10]. In the present paper we study the one-sided limit order book model of Lakner et al. [9], and our goal is to show how some properties of this model can be efficiently studied thanks to a coupling with a branching random walk.

From a high-level perspective, this coupling adds a new dimension to the initial limit order book model by creating a filiation between the orders. Such ideas have been extremely successful in queueing theory, see for instance Kendall [7], and we believe they can also be useful beyond the context of the model proposed here. For

E-mail address: f.simatos@tue.nl.

Date: October 29, 2012.

This work was done while the author was affiliated with CWI and sponsored by an NWO-VIDI grant.

instance, the model proposed by Yudovina [10] is also amenable to a tree representation, but the corresponding dynamic on trees is more challenging to analyze than the one here.

Branching random walks. The Galton Watson process is the simplest model of branching process. It represents the size of a population that evolves in discrete time, where at every time step each individual dies and is replaced by a random number of offspring, see for instance Athreya and Ney [1] for more details. A branching random walk is an extension of a Galton Watson process that adds a spatial component to the model. In addition to the genealogical structure given by the Galton Watson process, each individual has some location, say on the real line $\mathbb{R}$, that is given by a random displacement of her parent's location. Branching random walks can therefore be represented by trees with labels on the edges: the structure of the tree represents the genealogy of the underlying Galton Watson process, and the labels on the edges represent the displacement of the child with respect to her parent's location. In this paper we will consider the simplest model of branching random walks, where labels on the edges are i.i.d., and will use results by Biggins [2] and Biggins et al. [3] to study the limit order book model.

Acknowledgement. The author would like Josh Reed for introducing him to the limit order book model of the present paper, and Elie Aidekon for interesting discussions on branching random walks.

2. One-sided limit order book model

2.1. **Model.** Let us define a *book* as a finite point measure on $\mathbb{R}$ and an *order* as a point of a book. Let $\mathcal{B}$ be the set of books. For a book $\beta \in \mathcal{B}$, let $|\beta|$ be its mass (i.e., the number of orders it contains) and $\pi(\beta)$, which we call the *price* of the book, be the right endpoint of its support, i.e., the location of the rightmost order:

$$\pi(\beta) = \max\left\{x \in \mathbb{R} : \beta(\{x\}) > 0\right\}, \ \beta \in \mathcal{B}.$$

When the book is empty, i.e., $|\beta| = 0$, the value of the price is inconsequent for the purposes of this paper, say for instance $\pi(\beta) = 0$. Fix $p \in [0, 1]$ and a real-valued random variable X. We are interested in the $\mathcal{B}$-valued Markov chain $(B_n, n \geq 0)$ with the following dynamic.

If the book is empty, the process goes to state δ_0 in the next time step, where here and in the sequel δ_x stands for the Dirac mass at $x \in \mathbb{R}$. If the book is not empty, a coin with bias p is flipped. If heads (with probability p), an order is added to the current book at a random distance distributed according to X from the current price of the book, independently from everything else. If tails (thus, with probability $1 - p$), an order sitting at the current price is removed. Formally, the Markov chain (B_n) has the following dynamic: for any $\beta \in \mathcal{B}$ and any measurable function $f : \mathcal{B} \to [0, \infty)$,

$$\begin{aligned} (1) \quad \mathbb{E}\left[f(B_{n+1}) \mid B_n = \beta\right] = f(\delta_0)\mathbb{1}_{\{|\beta|=0\}} + p\mathbb{E}\left[f(\beta + \delta_{\pi(\beta)+X})\right]\mathbb{1}_{\{|\beta|>0\}} \\ + (1-p)f(\beta - \delta_{\pi(\beta)})\mathbb{1}_{\{|\beta|>0\}}. \end{aligned}$$

As explained in the introduction, this model (in continuous-time, and with a different boundary condition) has been proposed by Lakner et al. [9] to model a one-sided limit order book. The interpretation of the model is as follows: B_n

represents the state of a one-sided limit order book with only buy orders. In each time step, either a trader places a new buy order (with probability p), or a trader places a market order (to sell the commodity, with probability $1-p$). In the latter case, the trader sells the commodity at the highest available buy order, thus removing one order sitting at the price. From this perspective, the behavior of the price process $(\pi(B_n), n \geq 0)$ is of primary interest. In this paper, we show how one can answer questions related to the price process by coupling (B_n) with a branching random walk. Our coupling can be used to answer more elaborate questions on this particular model, and we believe that it also has the potential to be applied to other models.

Remark 1. To be more specific, in order to recover the model of Lakner et al. [9] one needs to apply an exponential transformation to (B_n), i.e., to consider the process $\widetilde{B}_n = \sum_{x \in B_n} \delta_{e^x}$. This transformation makes the orders live on $(0, \infty)$, in which case the interpretation of π as a price is reasonable. It also makes the price increase in a multiplicative rather than linear fashion, which is a common behavior in mathematical finance (e.g., geometric Brownian motion). We prefer to state the model on the line with linear displacement because of the analogy with branching random walks.

2.2. **Price dynamic.** The behavior of the price is asymmetric due to the system's dynamic. On the one hand, the price increases when an order is added to the right of the current price, and so an increase of the price is distributed according to X given that $X > 0$. On the other hand, the price decreases when an order is removed from the book, in which case the decrease of the price depends on the distance between the price and the second rightmost order. In particular, orders to the left of the price act as a barrier that slow down the price as it wants to drift downward.

Thus, although $\mathbb{E}X < 0$ seems at first a natural condition for the price to drift to $-\infty$, it seems plausible that if p is sufficiently close to 1, there will be so many orders sitting to the left of the price that they will eventually make the price drift to $+\infty$. This intuition turns out to be correct as Theorem 1 below shows.

This kind of behavior is strongly reminiscent of the behavior of extremal particles in branching random walks. There, although a typical particle drifts to $-\infty$ when $\mathbb{E}X < 0$, one may still observe atypical trajectories due to the exponential explosion in the number of particles, see the classical references by Hammersley [6], Kingman [8] and Biggins [2]. This analogy has actually been our initial motivation to investigate the relation between (B_n) and branching random walks. And indeed, we will show in Theorem 2 that B_n can be realized as some functional of a branching random walk, and this connection will make the proof of Theorem 1 quite transparent.

Theorem 1. *Assume that $p > 1/2$ and that $\mathbb{E}X$ exists in $(-\infty, \infty)$.*

If $\mathbb{E}X > 0$, then $\pi(B_n) \to +\infty$ almost surely.

Else, assume in addition to $p > 1/2$ that $\mathbb{E}X < 0$ and that $\mathbb{P}(X > 0) > 0$, and let $a = \inf_{\theta \geq 0} \mathbb{E}(e^{\theta X}) \in (0, 1]$. If $p > 1/(1+a)$, then $\pi(B_n) \to +\infty$ almost surely, while if $p < 1/(1+a)$ then $\pi(B_n) \to -\infty$ almost surely.

Remark 2. Let $\varphi(\theta) = \mathbb{E}(e^{\theta X})$ for $\theta \geq 0$ and assume that $\mathbb{E}X < 0$. We will use the following dichotomy: either $\varphi(\theta) = +\infty$ for every $\theta > 0$, in which case

$a = 1$; or $\varphi(\theta) < +\infty$ for some $\theta > 0$, in which case $a < 1$ due to the fact that $\varphi'(0) = \mathbb{E}X < 0$ in this case.

Note that the price process is recurrent if $p \leq 1/2$ since $(|B_n|, n \geq 0)$ is a random walk reflected at 0, so the previous result gives a full picture of the price behavior (except for the boundary cases $\mathbb{E}X = 0$ and $p = 1/(1+a)$). It is interesting to observe that if $p > 1/2$, $\mathbb{E}X < 0$, $\mathbb{P}(X > 0) > 0$ and X has a heavy right tail, in the sense that $\varphi(\theta) = +\infty$ for every $\theta > 0$ or equivalently, the random variable $\max(X, 0)$ has no finite exponential moment, then the price will always diverge to $+\infty$, irrespectively of the values of p and $\mathbb{E}X$. Although the fact that exponential moments play a key role is clear from a branching process perspective, we find it more surprising from the perspective of the limit order book.

3. Coupling with a branching random walk

3.1. **Intuition.** From the book process (B_n), one can construct a genealogical structure by making an order x a child of some other order y if x was added to the book at a time where y corresponded to the price of the book, i.e., x is added at a time n where $y = \pi(B_n)$. Since there is each time a probability p of adding an order to the book, it is intuitively clear that this construction will give each order (at most) a geometric number of offspring. By labeling the edge between x and y with the displacement $x - y$, which has distribution X, we end up with a Galton Watson tree with geometric offspring distribution and i.i.d. real-valued labels on the edges, i.e., a branching random walk. The idea of the coupling is to reverse this construction and to start from the branching random walk to build the book process (B_n). To do so, we will essentially realize the process (B_n) as the iteration of a deterministic tree operator Φ on a random tree, thus encoding all the randomness in the tree.

Nodes of the tree represent orders of the books, and in order to distinguish between orders that are currently in the book, orders that have been in the book and removed, and orders that have not been in the book so far (but may be later) we consider trees where nodes have one of three colors: green (orders currently in the book), red (orders removed from the book) and white (orders not added in the book so far). We also consider trees with real-valued labels on the edges: then, each node is also given a label by summing up the labels on the edges from the root to the node, the root having label 0. The label of a node represents the position of the corresponding order in the book. Then, the green node with largest label, say γ, represents the order at the current price, and so we will run the following dynamic on trees:

- if γ has at least one white child, then its first white child becomes green;
- if γ has no white child, then γ becomes red;
- if the tree has no green node then we need to draw a new random tree.

3.2. **The coupling: notation and main result.** Let $\mathcal{T}$ be the set of rooted trees where:

- every edge has a real-valued label;
- every node has one of three colors, green, red or white;
- finitely many nodes are green or red, and all green and red nodes are connected;
- the root is green or red.

We will need to compare the labels and colors of the nodes and edges of various trees. In that respect, it is convenient to consider $\mathcal{V}$ the set of all possible nodes and $\mathcal{E}$ the set of all possible edges, and to denote by $\mathcal{V}(t) \subset \mathcal{V}$ the set of nodes, $\mathcal{E}(t) \subset \mathcal{E}$ the set of edges and $\mathcal{G}(t), \mathcal{R}(t) \subset \mathcal{V}(t)$ the set of green and red nodes, respectively, of a tree $t \in \mathcal{T}$. Nodes inherit labels as explained above, i.e., the label of a node is obtained by summing up the labels on the edges from the root to this node, and the root has label 0. If $v \in \mathcal{V}$ is a node and $e \in \mathcal{E}$ is an edge, we denote by $\ell(v,t)$ and $\ell(e,t)$ the label of this node and edge in the tree $t \in \mathcal{T}$, provided $v \in \mathcal{V}(t)$ and $e \in \mathcal{E}(t)$. We call genealogical structure of a tree $t \in \mathcal{T}$ the tree obtained from t when forgetting about labels and colors, and we say that t is a subtree of t' and write $t \subset t'$ if the genealogical structure of t is a subtree of the genealogical structure of t' and $\ell(e,t) = \ell(e,t')$ for every $e \in \mathcal{E}(t) \subset \mathcal{E}(t')$. For $t \in \mathcal{T}$ let $\Gamma(t) \in \mathcal{B}$ be the point measure recording the labels of the green nodes of t:

$$\Gamma(t) = \sum_{v \in \mathcal{G}(t)} \delta_{\ell(v,t)}.$$

Let $\mathcal{T}^* = \{t \in \mathcal{T} : |\mathcal{G}(t)| > 0\}$ be the set of trees with at least one green node. If $t \in \mathcal{T}^*$ we denote by $\gamma(t)$ the green node with largest label and by $\omega(t)$ the number of white children of $\gamma(t)$. If there are several green nodes with maximal labels, we choose the first one (where in the sequel, nodes are ordered according to the lexicographical order). For $t \in \mathcal{T}$ and $v \in \mathcal{V}(t)$ we will more generally define $\omega(v,t)$ as the number of white children of v in t, so that $\omega(t) = \omega(\gamma(t),t)$ for $t \in \mathcal{T}^*$. The following operator will create the dynamic of (B_n).

Definition (Operator Φ)**.** Let $\Phi : \mathcal{T} \to \mathcal{T}$ be the following operator: if $|\mathcal{G}(t)| = 0$ then $\Phi(t) = t$, while if $t \in \mathcal{T}^*$, then Φ changes the color of one node according to the following rule:

- if $\omega(t) > 0$, Φ transforms the first white child of $\gamma(t)$ into a green node;
- if $\omega(t) = 0$, then Φ transforms $\gamma(t)$ into a red node.

We also define Φ_n as the nth iterate of Φ, defined by Φ_0 being the identity map and $\Phi_{n+1} = \Phi \circ \Phi_n$ for $n \geq 0$. Finally, we define $\kappa(t) = \inf\{n \geq 0 : |\mathcal{G}(\Phi_n(t))| = 0\}$ for $t \in \mathcal{T}$, so that $\kappa(t) \in \{0, 1, \ldots, \infty\}$, as the first time where iterating Φ on t creates a tree with no green node.

We can now state our main result.

Theorem 2. *Let T be the following random tree:*

- *the genealogical structure of T is a Galton Watson tree with geometric offspring distribution with parameter p;*
- *labels on the edges are i.i.d. with distribution X, independently from the genealogical structure;*
- *all nodes are white, except for the root which is green.*

Then $(\Gamma(\Phi_n(T)), 0 \leq n \leq \kappa(T))$ is equal in distribution to $(B_n, 0 \leq n \leq \tau)$ started at $B_0 = \delta_0$, where $\tau = \inf\{n \geq 0 : |B_n| = 0\}$.

3.3. **Proof of Theorem 1 based on Theorem 2.** Thanks to Theorem 2, we can write $B_n = \Gamma(\Phi_n(T))$ for $n \leq \kappa(T) = \tau$. The idea to prove Theorem 1 is, thanks to this representation, to translate the fact that the price drifts to $\pm\infty$ into conditions on the initial tree T, and then to analyze these conditions thanks to classical results from branching random walk theory.

In the rest of this section we assume that $p > 1/2$. Since we are interested in the long-time behavior of the price which goes back to 0 at time τ in the event $\{\tau < +\infty\}$, we consider everything in the event $\{\tau = \kappa(T) = +\infty\}$. It is not hard to show that $\kappa(T) = +\infty$ if and only if T is infinite, which happens with positive probability since we are in the supercritical case $p > 1/2$.

The case $\pi(B_n) \to -\infty$. Assume that $\mathbb{E}X < 0$, $\mathbb{P}(X > 0) > 0$ and $p < 1/(1+a)$: we want to prove that $\pi(B_n) \to -\infty$. Since $p > 1/2$ by assumption, we have in particular $a < 1$ and so there must exist $\eta > 0$ such that $\mathbb{E}(e^{\eta X}) < +\infty$ (see Remark 2 following Theorem 1).

Let M_n be the rightmost point of the branching random walk T at time n, i.e.,

$$M_n = \max\{\ell(v, T) : v \in \mathcal{V}(T) \text{ and } |v| = n\}$$

where $|v|$ is the distance from v to the root. Under the assumptions made on X and p, Theorem 4 in Biggins [2] shows that $M_n \to -\infty$ almost surely (in the event of non-extinction). Thus to conclude the proof of $\pi(B_n) \to -\infty$, we only have to show that $M_n \to -\infty$ implies that $\pi(B_n) \to -\infty$.

Let $K \geq 0$ and n_0 such that $M_n \leq -K$ for any $n \geq n_0$, i.e., $\ell(v, T) \leq -K$ for every $v \in \mathcal{V}(T)$ with $|v| \geq n_0$. Since there are only finite many nodes of T at depth $< n_0$, we must have $|\gamma(\Phi_n(T))| \geq n_0$ for n large enough and for those n, we have $\ell(\gamma(\Phi_n(T)), \Phi_n(T)) \leq -K$ by choice of n_0. This proves that $\pi(B_n) \to -\infty$.

The case $\pi(B_n) \to +\infty$. Let q be the probability that the price remains non-negative at all times: $q = \mathbb{P}(\pi(B_n) \geq 0, \forall n)$. Observe that when a new order is placed in the book to the right of the current price, say at time n and at $x > \pi(B_n)$, this order defines the new price, and the probability that the price never falls back below x after time n is exactly q. Thus, if $q > 0$ we can define a renewal sequence corresponding to the times at which a new order is added to the book, which defines a level above which the price stays forever. Then the renewal theorem implies that $\pi(B_n) \to +\infty$.

Thus, $q > 0$ is a sufficient condition for $\pi(B_n) \to +\infty$. But to compute q, one may as well assume that orders placed on $(-\infty, 0)$ are instantaneously removed since these orders only play a role when the price visits $(-\infty, 0)$. In terms of T, this amounts to removing all nodes with negative labels, as well as all their descendants. In more geometrical terms, this amounts to set a barrier at 0 and kill every individual (and hence its descendants) of the branching random walk that crosses this barrier, which is precisely the model studied in Biggins et al. [3]. In this context, q is the probability that this branching random walk with a barrier at 0 survives for ever.

When $\mathbb{E}X > 0$, it is clear that $q > 0$: indeed, any line of descent of T is a random walk with positive drift, which thus has positive probability of never visiting $(-\infty, 0)$.

When $\mathbb{E}X < 0$ and $\mathbb{P}(X > 0) > 0$, the situation is more delicate, but Theorem 1 in Biggins et al. [3] asserts that $q > 0$ if $p > 1/(1+a)$ and $\mathbb{E}(e^{\eta X}) < +\infty$ for some $\eta > 0$. Thus, to complete the proof of Theorem 1, it remains only to prove that $q > 0$ if $\mathbb{E}X < 0$, $\mathbb{P}(X > 0) > 0$ and $\mathbb{E}(e^{\eta X}) = +\infty$ for every $\eta > 0$ (note that in this case, we are always in the case $p > 1/(1+a)$ since $a = 1$, see Remark 2). We prove this by a truncation argument. Let $X_K = \min(X, K)$ and $a_K = \inf_\theta \mathbb{E}(e^{\theta X_K})$: since $a = 1$ we have $a_K \to 1$ as $K \to +\infty$, and hence $p > 1/(1+a_K)$ for K large

enough. For such a K, we have $q_K > 0$, with q_K the same probability q as above but when the displacements have distribution X_K instead of X. By monotonicity (which is obvious on the branching random walk with a barrier), this proves the desired result $q > 0$ and finally concludes the proof of Theorem 1.

4. Proof of Theorem 2

4.1. **Study of an auxiliary tree-valued Markov chain.** Theorem 2 is very intuitive. Unfortunately, the rigorous proof involves quite a lot of formalism, since we need to go into the details of the tree dynamic induced by iterations of Φ. In order to slightly reduce the notational burden, we will assume that X is a discrete random variable; it is just a matter of formalism to extend the proof below to the general case. Let us introduce the following tree operators:

- for $t \in \mathcal{T}$, $\Upsilon(t)$ is the tree obtained from t by deleting all white nodes;
- for $t \in \mathcal{T}_\Omega = \{t \in \mathcal{T} : |\mathcal{G}(t)| > 0 \text{ and } \omega(t) > 0\}$, $\Omega(t)$ is the tree obtained from t by turning the first white child of $\gamma(t)$ into a green node;
- for $t \in \mathcal{T}^*$ and $\lambda \in \mathbb{R}$, $\Omega'(t, \lambda)$ is the tree obtained from t by adding a green child to $\gamma(t)$ with label λ on the corresponding edge;
- for $t \in \mathcal{T}^*$, $\Psi(t)$ is the tree obtained from t by turning $\gamma(t)$ into a red node.

It will also be convenient to introduce the following subsets of $\mathcal{T}$:

- $\mathcal{T}_\Upsilon$ is the set of finite rooted trees with only green or red nodes;
- $\mathcal{T}_\Psi = \{t \in \mathcal{T} : |\mathcal{G}(t)| > 0 \text{ and } \omega(t) = 0\}$;
- $\mathcal{T}_0$ is the set of trees of which every node is white, except for the root which is green.

Note that $\mathcal{T} = \{t : |\mathcal{G}(t)| = 0\} \cup \mathcal{T}_\Omega \cup \mathcal{T}_\Psi$. Let $Y_n = \Upsilon(\Phi_n(T))$ with $T \in \mathcal{T}_0$ as in Theorem 2. We denote by $t_0 = \Upsilon(T)$ the deterministic tree reduced to the root, which is green. The goal of this section is to prove that the process $(Y_n, n \geq 0)$ defines a Markov chain started at t_0 with the following dynamic: for any $n \geq 0$, any $y, y' \in \mathcal{T}_\Upsilon$ (note that for any $t \in \mathcal{T}_0$ and any $k \geq 0$, $\Upsilon(\Phi_k(t)) \in \mathcal{T}_\Upsilon$) and any $x \in \mathbb{R}$,

$$\mathbb{P}(Y_{n+1} = y' \mid Y_n = y) = \begin{cases} 1 & \text{if } |\mathcal{G}(y)| = 0 \text{ and } y' = y, \\ p\mathbb{P}(X = x) & \text{if } |\mathcal{G}(y)| > 0 \text{ and } y' = \Omega'(y, x), \\ 1 - p & \text{if } |\mathcal{G}(y)| > 0 \text{ and } y' = \Psi(y), \\ 0 & \text{otherwise.} \end{cases}$$

To this end, we fix until the rest of this section $n \geq 0$, $x \in \mathbb{R}$ and $y_k \in \mathcal{T}_\Upsilon$ for $k = 0, \ldots, n+1$, and we aim to prove that

$$\text{(2)} \quad \mathbb{P}\left(Y_k = y_k, 0 \leq k \leq n+1\right) =$$
$$\mathbb{P}\left(Y_k = y_k, 0 \leq k \leq n\right) \times \begin{cases} 1 & \text{if } |\mathcal{G}(y_n)| = 0 \text{ and } y_{n+1} = y_n, \\ p\mathbb{P}(X = x) & \text{if } |\mathcal{G}(y_n)| > 0 \text{ and } y_{n+1} = \Omega'(y_n, x), \\ 1 - p & \text{if } |\mathcal{G}(y_n)| > 0 \text{ and } y_{n+1} = \Psi(y_n) \end{cases}$$

which will prove the Markov property of (Y_n) with the prescribed dynamic. Let us study the dynamic of (Y_n). According to the various definitions made, we have for any $t \in \mathcal{T}$ that $\Phi(t)$ is equal to t if $|\mathcal{G}(t)| = 0$, to $\Omega(t)$ if $t \in \mathcal{T}_\Omega$ and to $\Psi(t)$ if

$t \in \mathcal{T}_\Psi$, so that

$$\Upsilon(\Phi(t)) = \begin{cases} \Upsilon(t) & \text{if } |\mathcal{G}(t)| = 0, \\ \Upsilon(\Omega(t)) & \text{if } t \in \mathcal{T}_\Omega, \\ \Upsilon(\Psi(t)) & \text{if } t \in \mathcal{T}_\Psi. \end{cases}$$

It is clear that if $t \in \mathcal{T}_\Omega$, then $\Upsilon(\Omega(t)) = \Omega'(\Upsilon(t), \ell(t))$ with $\ell(t)$ the label on the edge between $\gamma(t)$ and its first white child, while if $t \in \mathcal{T}_\Psi$, then $\Upsilon(\Psi(t)) = \Psi(\Upsilon(t))$. Thus the previous display can be rewritten as

$$\Upsilon(\Phi(t)) = \begin{cases} \Upsilon(t) & \text{if } |\mathcal{G}(t)| = 0, \\ \Omega'(\Upsilon(t), \ell(t)) & \text{if } t \in \mathcal{T}_\Omega, \\ \Psi(\Upsilon(t)) & \text{if } t \in \mathcal{T}_\Psi \end{cases}$$

and since $\Phi_{n+1}(t) = \Phi(\Phi_n(t))$ we get

$$\Upsilon(\Phi_{n+1}(t)) = \begin{cases} \Upsilon(\Phi_n(t)) & \text{if } |\mathcal{G}(\Phi_n(t))| = 0, \\ \Omega'(\Upsilon(\Phi_n(t)), \ell(\Phi_n(t))) & \text{if } \Phi_n(t) \in \mathcal{T}_\Omega, \\ \Psi(\Upsilon(\Phi_n(t))) & \text{if } \Phi_n(t) \in \mathcal{T}_\Psi. \end{cases}$$

Since Υ does not affect the colors of green nodes, we have $\mathcal{G}(t) = \mathcal{G}(\Upsilon(t))$ and in particular, $|\mathcal{G}(\Phi_n(t))| = |\mathcal{G}(\Upsilon(\Phi_n(t)))|$. Plugging in the definitions of $\mathcal{T}_\Omega$ and $\mathcal{T}_\Psi$, it follows that for any $n \geq 0$ and any $t \in \mathcal{T}_0$, we have

$$(3) \quad \Upsilon(\Phi_{n+1}(t)) = \\ \begin{cases} \Upsilon(\Phi_n(t)) & \text{if } |\mathcal{G}(\Upsilon(\Phi_n(t)))| = 0, \\ \Omega'(\Upsilon(\Phi_n(t)), \ell(\Phi_n(t))) & \text{if } |\mathcal{G}(\Upsilon(\Phi_n(t)))| > 0 \text{ and } \omega(\Phi_n(t)) > 0, \\ \Psi(\Upsilon(\Phi_n(t))) & \text{if } |\mathcal{G}(\Upsilon(\Phi_n(t)))| > 0 \text{ and } \omega(\Phi_n(t)) = 0. \end{cases}$$

This last equation shows that $\Upsilon(\Phi_{n+1}(t))$ is almost entirely determined by $\Upsilon(\Phi_n(t))$, up to the knowledge (hidden by the action of Υ) of whether $\gamma(\Phi_n(t))$ has at least one white child in $\Phi_n(t)$ or not, and the value $\ell(\Phi_n(t))$ of the corresponding edge. Further, since $Y_n = \Upsilon(\Phi_n(T))$, (3) leads to

$$\mathbb{P}(Y_k = y_k, 0 \leq k \leq n+1) = \mathbb{1}_{\{|\mathcal{G}(y_n)|=0, y_{n+1}=y_n\}} \mathbb{P}(Y_k = y_k, 0 \leq k \leq n) \\ + \mathbb{1}_{\{|\mathcal{G}(y_n)|>0, y_{n+1}=\Omega'(y_n,x)\}} \mathbb{P}(\omega(\Phi_n(T)) > 0, \ell(\Phi_n(T)) = x, Y_k = y_k, 0 \leq k \leq n) \\ + \mathbb{1}_{\{|\mathcal{G}(y_n)|>0, y_{n+1}=\Psi(y_n)\}} \mathbb{P}(\omega(\Phi_n(T)) = 0, Y_k = y_k, 0 \leq k \leq n)$$

and so to prove (2), we only have to show that if $|\mathcal{G}(y_n)| > 0$, then

$$(4) \quad \mathbb{P}(\omega(\Phi_n(T)) > 0, \ell(\Phi_n(T)) = x, Y_k = y_k, 0 \leq k \leq n) \\ = p\mathbb{P}(X = x)\mathbb{P}(Y_k = y_k, 0 \leq k \leq n).$$

This property is quite intuitive: given the history Y_k for $k \leq n$ does not give any information on the remaining number of white children of $\gamma(\Phi_n(T))$ in $\Phi_n(T)$, nor on the label on the edge between $\gamma(\Phi_n(T))$ and its first white child, if any. The fact that every node has a geometric number of offspring and that labels on the edges are i.i.d. should therefore imply (4). To formalize this intuition, we will prove the following result, from which one can readily deduce (4). For $t \in \mathcal{T}$, let in the rest of the paper $\eta(v, t)$ be the number of children of the node $v \in \mathcal{V}(t)$.

Proposition. *If $|\mathcal{G}(y_n)| > 0$ and $\mathbb{P}(Y_k = y_k, 0 \le k \le n) > 0$, then for any $t \in \mathcal{T}_0$ we have*

$$\Upsilon(\Phi_k(t)) = y_k, 0 \le k \le n \iff y_n \subset t \text{ and } \eta(v, y_n) = \eta(v,t) \text{ for every } v \in \mathcal{R}(y_n).$$

Proof. For $t \in \mathcal{T}$, let in the rest of the proof $\sigma(t) = |\mathcal{G}(t)| + 2|\mathcal{R}(t)| - 1$, and recall that $\kappa(t) = \inf\{n \ge 0 : |\mathcal{G}(\Phi_n(t))| = 0\}$. It is clear from the definition of Φ that $\sigma(\Phi(t)) = \sigma(t) + \mathbb{1}_{\{|\mathcal{G}(t)|>0\}}$. Since $\Phi_n(t) = t$ for any $n \ge 0$ if $|\mathcal{G}(t)| = 0$ and $\sigma(t) = 0$ for $t \in \mathcal{T}_0$, it follows that

$$\forall t \in \mathcal{T}_0,\ \sigma(\Upsilon(\Phi_k(t))) = k \iff k \le \kappa(t). \tag{5}$$

We break the proof of the proposition into two steps.

First step. Fix some $t \in \mathcal{T}_0$ and $y \in \mathcal{T}_\Upsilon$. The first step of the proof consists in proving that the following conditions are equivalent:

(i) $\Upsilon(\Phi_{\sigma(y)}(t)) = y$;
(ii) there exists $t' \in \mathcal{T}_0$ such that $\Upsilon(\Phi_{\sigma(y)}(t')) = y$, $y \subset t$ and $\eta(v,y) = \eta(v,t)$ for every $v \in \mathcal{R}(y)$.

Proof of (i)$\Rightarrow$(ii). Assume that (i) holds, i.e., $\Upsilon(\Phi_{\sigma(y)}(t)) = y$: we want to prove (ii). Then taking $t' = t$ gives the existence of the desired t'. Moreover, since $\Phi(a)$ does not change the genealogical structure of $a \in \mathcal{T}$ and $\Upsilon(a)$ only truncates a, we have $\Upsilon(\Phi_n(t)) \subset t$ for any $n \ge 0$, in particular $y \subset t$. Then, consider any $v \in \mathcal{R}(y)$. Since all the nodes of t except for the root are white, the color of v results from the successive applications of Φ to $t \in \mathcal{T}_0$. In particular, v being red in $\Phi_{\sigma(y)}(t)$ comes from the fact that at some point, none of the children of v were white, i.e., $\omega(v, \Phi_k(t)) = 0$ for some $k \le \sigma(y)$. Since Φ does not create white nodes, this implies $\omega(v, \Phi_{\sigma(y)}(t)) = 0$ and since Υ conserves all non-white nodes, v has as many children in $\Phi_{\sigma(y)}(t)$ as in $\Upsilon(\Phi_{\sigma(y)}(t)) = y$, i.e., $\eta(v, \Phi_{\sigma(y)}(t)) = \eta(v,y)$ which gives $\eta(v,t) = \eta(v,y)$.

Proof of (ii)$\Rightarrow$(i). Assume that (ii) holds: we want to prove (i). So in the rest of the proof, consider some $t' \in \mathcal{T}_0$ such that $\Upsilon(\Phi_{\sigma(y)}(t')) = y$, and assume that $y \subset t$ and $\eta(v,y) = \eta(v,t)$ for every $v \in \mathcal{R}(y)$. We prove that $y = \Upsilon(\Phi_{\sigma(y)}(t))$ by induction on $\sigma(y)$.

If $\sigma(y) = 0$, then on the one hand, $\Upsilon(\Phi_{\sigma(y)}(t)) = t_0$ while on the other hand, $|\mathcal{G}(y)| + 2|\mathcal{R}(y)| - 1 = 0$ implies $|\mathcal{G}(y)| = 1$ and $|\mathcal{R}(y)| = 0$, so that $y = t_0$. Thus $y = \Upsilon(\Phi_{\sigma(y)}(t))$ when $\sigma(y) = 0$, which initializes the induction.

Assume now that $\sigma(y) \ge 1$. Then $\sigma(y) = \sigma(\Upsilon(\Phi_{\sigma(y)}(t'))$ and so (5) implies that $\sigma(y) \le \kappa(t')$. Define $y' = \Upsilon(\Phi_{\sigma(y)-1}(t'))$: since $\sigma(y) - 1 \le \kappa(t')$, (5) implies that $\sigma(y') = \sigma(y) - 1$, i.e., $\sigma(\Phi_{\sigma(y)}(t')) = 1 + \sigma(\Phi_{\sigma(y)-1}(t'))$. This last equality means that applying Φ on $\Phi_{\sigma(y)-1}(t')$ creates a green node or changes a green node into a red one, meaning in every case that $|\mathcal{G}(\Phi_{\sigma(y)-1}(t'))| > 0$ and since $\mathcal{G}(\Phi_{\sigma(y)-1}(t')) = \mathcal{G}(\Upsilon(\Phi_{\sigma(y)-1}(t'))) = \mathcal{G}(y')$ this finally means that $|\mathcal{G}(y')| > 0$. In view of $y = \Upsilon(\Phi_{\sigma(y)}(t'))$, $y' = \Upsilon(\Phi_{\sigma(y)-1}(t'))$ and (3), we therefore only have two possibilities:

$$y = \begin{cases} \Omega'(y', \ell(\Phi_{\sigma(y)-1}(t'))) & \text{if } \omega(\Phi_{\sigma(y)-1}(t')) > 0, \\ \Psi(y') & \text{if } \omega(\Phi_{\sigma(y)-1}(t')) = 0. \end{cases} \tag{6}$$

In either case, we have $y' \subset y$ and since $y \subset t$ by assumption, this gives $y' \subset t$. Moreover, the action of Ω' is to add one green node to γ and the action of Ψ is to turn γ into a red node, so that in either case we have $\mathcal{R}(y') \subset \mathcal{R}(y)$ and $\eta(v, y') = \eta(v, y)$ for every $v \in \mathcal{R}(y')$. Since $\eta(v, y) = \eta(v, t)$ for every $v \in \mathcal{R}(y)$ by assumption, this implies that $\eta(v, y') = \eta(v, t)$ for every $v \in \mathcal{R}(y')$. Since finally $y' = \Upsilon(\Phi_{\sigma(y')}(t'))$ and $\sigma(y') = \sigma(y) - 1 < \sigma(y)$, we can therefore invoke the induction hypothesis to deduce that $\Upsilon(\Phi_{\sigma(y')}(t)) = y' = \Upsilon(\Phi_{\sigma(y')}(t'))$. In particular, $\Phi_{\sigma(y')}(t)$ and $\Phi_{\sigma(y')}(t')$ have the same set of green and red nodes and since $|\mathcal{G}(\Phi_{\sigma(y')}(t'))| > 0$ this shows that $|\mathcal{G}(\Phi_{\sigma(y')}(t))| > 0$. Then, (3) shows that

$$\Upsilon(\Phi_{\sigma(y)}(t)) = \begin{cases} \Omega'(\Upsilon(\Phi_{\sigma(y')}(t)), \ell(\Phi_{\sigma(y')}(t))) & \text{if } \omega(\Phi_{\sigma(y')}(t)) > 0, \\ \Psi(\Upsilon(\Phi_{\sigma(y')}(t))) & \text{if } \omega(\Phi_{\sigma(y')}(t)) = 0. \end{cases}$$

Since $y' = \Upsilon(\Phi_{\sigma(y')}(t))$ this can be rewritten as

$$\Upsilon(\Phi_{\sigma(y)}(t)) = \begin{cases} \Omega'(y', \ell(\Phi_{\sigma(y')}(t))) & \text{if } \omega(\Phi_{\sigma(y')}(t)) > 0, \\ \Psi(y') & \text{if } \omega(\Phi_{\sigma(y')}(t)) = 0, \end{cases}$$

and in view of (6), the proof of $y = \Upsilon(\Phi_{\sigma(y)}(t))$ will be complete if we can show the two following implications:

$$\omega(\Phi_{\sigma(y')}(t')) > 0 \Rightarrow \omega(\Phi_{\sigma(y')}(t)) > 0 \text{ and } \ell(\Phi_{\sigma(y')}(t)) = \ell(\Phi_{\sigma(y')}(t'))$$

and

$$\omega(\Phi_{\sigma(y')}(t')) = 0 \Rightarrow \omega(\Phi_{\sigma(y')}(t)) = 0.$$

To prove these two implications, we will use the identities

$$\omega(\Phi_{\sigma(y')}(t')) = \eta(\gamma(y'), t') - \eta(\gamma(y'), y') \tag{7}$$

and

$$\omega(\Phi_{\sigma(y')}(t)) = \eta(\gamma(y'), t) - \eta(\gamma(y'), y') \tag{8}$$

that come from the fact that $\Upsilon(\Phi_{\sigma(y')}(t')) = \Upsilon(\Phi_{\sigma(y')}(t))$.

- **Direct implication:** assume that $\omega(\Phi_{\sigma(y')}(t')) > 0$. In this case, applying Φ to $\Phi_{\sigma(y')}(t')$ adds a green child to $\gamma(y')$ in $\Phi_{\sigma(y)}(t')$, and so $\eta(\gamma(y'), y) = 1 + \eta(\gamma(y'), y')$. Since $y \subset t$ by assumption, this gives $\eta(\gamma(y'), t) > \eta(\gamma(y'), y')$ which proves that $\omega(\Phi_{\sigma(y')}(t)) > 0$ in view of (7). Moreover, the equality $\eta(\gamma(y'), \Upsilon(\Phi_{\sigma(y')}(t'))) = \eta(\gamma(y'), \Upsilon(\Phi_{\sigma(y')}(t)))$ means that $\gamma(y')$ has as many green and red children in $\Phi_{\sigma(y')}(t')$ than in $\Phi_{\sigma(y')}(t)$. In particular, applying Φ to these two trees adds the same node, say v, to each tree. Then, the equality $\ell(\Phi_{\sigma(y')}(t)) = \ell(\Phi_{\sigma(y')}(t'))$ comes from the fact that the label on the edge between $\gamma(y')$ and v is the same in t and t' due to the inclusion $y = \Upsilon(\Phi_{\sigma(y)}(t')) \subset t$;
- **Reverse implication:** assume that $\omega(\Phi_{\sigma(y')}(t')) = 0$. In this case, applying Φ to $\Phi_{\sigma(y')}(t')$ turns $\gamma(y')$ into a red node in $\Phi_{\sigma(y)}(t')$, i.e., $\gamma(y') \in \mathcal{R}(y)$ and in particular, $\eta(\gamma(y'), y) = \eta(\gamma(y'), t)$ by assumption on y. On the other hand, Φ did not change the genealogical structure of y' and so $\eta(\gamma(y'), y) = \eta(\gamma(y'), y')$ which shows that $\eta(\gamma(y'), y') = \eta(\gamma(y'), t)$. This implies $\omega(\Phi_{\sigma(y')}(t)) = 0$ in view of (8).

This concludes the proof of the first step.

Second step. We now prove the proposition. Assume in the rest of the proof that $|\mathcal{G}(y_n)| > 0$ and $\mathbb{P}\left(Y_k = y_k, 0 \le k \le n\right) > 0$. The direct implication is rather straightforward: if $\Upsilon(\Phi_k(t)) = y_k$ for every $k = 0, \ldots, n$, then for $k = n$ the result of the first step implies that $y_n \subset t$ and $\eta(v, y_n) = \eta(v, t)$ for every $v \in \mathcal{R}(y_n)$.

Let us now prove the converse implication, so assume that $y_n \subset t$ and $\eta(v, y_n) = \eta(v, t)$ for every $v \in \mathcal{R}(y_n)$. The goal is to prove that $\Upsilon(\Phi_k(t)) = y_k$ for every $k = 0, \ldots, n$. Since $\mathbb{P}\left(Y_k = y_k, 0 \le k \le n\right) > 0$, there exists $t' \in \mathcal{T}_0$ such that $y_k = \Upsilon(\Phi_k(t'))$ for every $k = 0, \ldots, n$. Because Φ does not erase nodes and never changes the color of a red node, we have $y_k \subset y_{k+1}$ and $\mathcal{R}(y_k) \subset \mathcal{R}(y_{k+1})$ for any $0 \le k \le n-1$. In particular, for every $0 \le k \le n$ it holds that $y_k \subset t$ and $\eta(v, y_k) = \eta(v, t)$ for every $v \in \mathcal{R}(y_k)$. Moreover, Φ cannot create green nodes starting from a tree with no green node, and so the condition $|\mathcal{G}(y_n)| > 0$ implies that $|\mathcal{G}(y_k)| > 0$ for every $k \le n$. Thus $|\mathcal{G}(\Phi_k(t'))| > 0$ and so $k < \kappa(t)$, so that (5) implies that $\sigma(y_k) = k$, i.e., $y_k = \Upsilon(\Phi_{\sigma(y_k)}(t'))$. Thus all the assumptions of the first step are satisfied, and we deduce that $y_k = \Upsilon(\Phi_{\sigma(y_k)}(t)) = \Upsilon(\Phi_k(t))$ for every $k \le n$. This finally concludes the proof of the proposition. $\square$

4.2. **Proof of Theorem 2.** By regeneration of (B_n) at $\tau + 1$ and in view of (1), it is enough to show that $(\Gamma(\Phi_n(T)), n \ge 0)$ is the Markov chain started at δ_0 with the following transition: for any $\beta \in \mathcal{B}$ and any measurable function $f : \mathcal{B} : [0, \infty)$,

$$\begin{aligned}\mathbb{E}\left[f\big(\Gamma(\Phi_{n+1}(T))\big) \mid \Gamma(\Phi_n(T)) = \beta\right] = f(\delta_0)\mathbb{1}_{\{|\beta|=0\}} &+ p\mathbb{E}\left[f(\beta + \delta_{\pi(\beta)+X})\right]\mathbb{1}_{\{|\beta|>0\}}\\ &+ (1-p) f(\beta - \delta_{\pi(\beta)})\mathbb{1}_{\{|\beta|>0\}}.\end{aligned}$$

Since Γ only depends on the green nodes, we have $\Gamma(\Phi_n(T)) = \Gamma(Y_n)$ and so we only have to show that $\Gamma(Y_n)$ is the Markov chain with the prescribed dynamic. This is easily verified once one realizes that the σ-algebras $\sigma(Y_k, 0 \le k \le n)$ and $\sigma(\Gamma(Y_k), 0 \le k \le n)$ are the same. Indeed, the inclusion $\sigma(\Gamma(Y_k), 0 \le k \le n) \subset \sigma(Y_k, 0 \le k \le n)$ is trivial. For the reverse inclusion, note that from the sequence $(\Gamma(Y_k), 0 \le k \le n)$ one can recover the sequence $(Y_k, 0 \le k \le n)$. This can be proved by induction on n. For $n = 0$ this is trivial, since $Y_0 = t_0$ and $\Gamma(Y_0) = \delta_0$. So assume it holds for $n \ge 0$, and let us prove it for $n + 1$. So assume that $(\Gamma(Y_k), 0 \le k \le n+1)$ is known. Then by induction hypothesis, $(Y_k, 0 \le k \le n)$ is known. Moreover, (3) shows that there are only three possible cases:

- either $Y_{n+1} = Y_n$, in which case $\Gamma(Y_{n+1}) = \Gamma(Y_n)$;
- or $Y_{n+1} = \Omega'(Y_n, \lambda)$ for some $\lambda \in \mathbb{R}$: in this case, $\Gamma(Y_{n+1}) = \Gamma(Y_n) + \delta_\lambda$;
- or $Y_{n+1} = \Psi(Y_n)$, in which case $\Gamma(Y_{n+1}) = \Gamma(Y_n) - \delta_\lambda$ for some $\lambda \in \mathbb{R}$.

Thus by comparing $\Gamma(Y_{n+1})$ to $\Gamma(Y_n)$ one can recover Y_{n+1} which shows that $\sigma(\Gamma(Y_k), 0 \le k \le n) \subset \sigma(Y_k, 0 \le k \le n)$. From this, one easily deduces using the Markov property of (Y_n) that $(\Gamma(Y_n))$ is a Markov chain with the prescribed dynamic. The proof of Theorem 2 is therefore complete.

References

[1] Athreya, K. B. and Ney, P. E. (2004). *Branching processes.* Dover Publications Inc., Mineola, NY. Reprint of the 1972 original [Springer, New York; MR0373040].

[2] Biggins, J. D. (1976). The first- and last-birth problems for a multitype age-dependent branching process. *Advances in Appl. Probability* **8,** 446–459.

[3] Biggins, J. D., Lubachevsky, B. D., Shwartz, A. and Weiss, A. (1991). A branching random walk with a barrier. *Ann. Appl. Probab.* **1,** 573–581.

[4] Cont, R., Stoikov, S. and Talreja, R. (2010). A stochastic model for order book dynamics. *Oper. Res.* **58,** 549–563.

[5] Gould, M. D., Porter, M. A., Williams, S., McDonald, M., Fenn, D. and Howison, S. Limit order books. Available at SSRN: http://ssrn.com/abstract=1970185, April 2012.

[6] Hammersley, J. M. (1974). Postulates for subadditive processes. *Ann. Probability* **2,** 652–680.

[7] Kendall, D. G. (1951). Some problems in the theory of queues. *Journal of the Royal Statistical Society. Series B (Methodological)* **13,** 151–185.

[8] Kingman, J. F. C. (1975). The first birth problem for an age-dependent branching process. *Ann. Probability* **3,** 790–801.

[9] Lakner, P., Reed, J. and Stoikov, S. A model for the limit order book. In preparation.

[10] Yudovina, E. A simple model of a limit order book. Submitted, 2012. arXiv:1205.7017.